Bayesian Seismo-Acoustic Inversion to Investigate Spatial Variability and Uncertainty of Shallow Water Sediments

Jan Dettmer
School of Earth and Ocean Sciences, University of Victoria, Victoria BC Canada
phone: (250) 472-4342 email: jand@uvic.ca

Stan E. Dosso
School of Earth and Ocean Sciences, University of Victoria, Victoria BC Canada
Tel: (250) 472-4341 email: sdosso@uvic.ca

Charles W. Holland
Pennsylvania State University, Applied Research Laboratory, State College PA
Tel: (814) 865-1724 email: holland-cw@psu.edu

Award Number: N000140710540

LONG-TERM GOALS

Propagation and reverberation of acoustic fields in shallow waters depend strongly on the spatial variability of seabed geoacoustic parameters, and lack of knowledge of seabed variability is often a limiting factor in acoustic modeling applications. However, direct sampling (e.g., coring) of vertical and lateral variability is expensive and laborious, and matched-field and other long-range inversion methods fail to provide sufficient resolution. The long-term goal of this work is to use a Bayesian inversion approach in combination with seabed reflectivity data to investigate and quantify spatial variability of seabed sediments. For proper quantitative examination of spatial variability, it is important to differentiate between parameter estimate uncertainty, model parameterization effects and actual spatial variability.

To date, the postdoctoral project has further developed and applied a probabilistic 1-D inversion method to quantify uncertainty of geoacoustic seabed parameters. In particular, rigorous methods for selecting the optimal model parameterizations (e.g., the number of sediment layers) were examined and applied. The ultimate goal of quantifying spatial variability is currently being pursued and will be continued in the remaining time of the project. Inversion results for multiple locations will be analyzed and compared to understand sediment variability.

OBJECTIVES

The objective of the research is to extend and use the inversion tools developed in Dettmer's PhD project to investigate geoacoustic variability and the associated uncertainties at fine spatial scales. Spatial variability will be investigated by interpreting the differences between the recovered geoacoustic profiles while taking quantitative parameter uncertainties into account. Quantitative uncertainties depend on the model parameterizations selected for the problem. Therefore, model parametrization needs to be selected according to a quantitative criterion. The small seafloor footprint and high

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
30 SEP 2008	**Annual**	**00-00-2008 to 00-00-2008**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Bayesian Seismo-Acoustic Inversion To Investigate Spatial Variability And Uncertainty Of Shallow Water Sediments	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Victoria,School of Earth and Ocean Sciences,Victoria BC Canada,**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
code 1 only

14. ABSTRACT

Propagation and reverberation of acoustic fields in shallow waters depend strongly on the spatial variability of seabed geoacoustic parameters, and lack of knowledge of seabed variability is often a limiting factor in acoustic modeling applications. However, direct sampling (e.g., coring) of vertical and lateral variability is expensive and laborious, and matched-field and other long-range inversion methods fail to provide sufficient resolution. The long-term goal of this work is to use a Bayesian inversion approach in combination with seabed reflectivity data to investigate and quantify spatial variability of seabed sediments. For proper quantitative examination of spatial variability, it is important to differentiate between parameter estimate uncertainty, model parameterization effects and actual spatial variability

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**9**	

precision of the reflectivity method provide the high lateral and vertical resolution required for this approach. The computation of quantitative uncertainty estimates for the recovered profiles is of key importance to differentiate between spatial variability of environmental parameters and uncertainties inherent in the inversion results.

APPROACH

This work uses Bayesian inference to determine model parameters of sediment profiles and their uncertainties from seismo-acoustic single bounce reflectivity measurements with small seafloor footprints ($\sim$100 m). Lateral variability can be inferred from differences in inversion results for appropriately spaced measurement sites.

Bayesian inversion formulates an inverse problem in terms of the posterior probability density (PPD) of the model parameters, incorporating both data and prior information. The solution is typically quantified in terms of properties of the multi-dimensional PPD representing parameter estimates, uncertainties and inter-relationships. Optimal parameter estimates require nonlinear optimization such as adaptive hybrid inversion (Dosso 2002). Parameter uncertainties (e.g., marginal distributions, credibility intervals) are computed using Markov-chain Monte Carlo methods (Dosso 2002; Dettmer et al. 2008a).

Rigorous uncertainty estimation for geoacoustic parameters is of key importance to meaningfully resolve spatial variability between inversion results for nearby measurement sites from the inherent inversion uncertainties (an ultimate goal of our work). This requires a nonlinear inversion approach, rigorous estimation of the data error statistics, and quantitative model selection.

Uncertainty estimates also depend on the model parametrization chosen for the inversion. Bayesian evidence is the basis for this model selection. Evidence brings a natural parsimony to the model selection problem which is referred to as the Bayesian razor. Estimating evidence is challenging due to the requirement to integrate the likelihood with respect to the prior (Chib 1995), and finding robust and accurate estimators for the evidence integral has seen the attention of much research. Due to the high computational demands of the forward and inverse problems considered in this paper, an asymptotic point estimate (for the MAP model vector) is used to carry out model selection. The Bayesian information criterion (BIC, Schwartz (1978)) is an asymptotic approximation derived for diffuse multivariate normal prior distributions (Kass and Raftery 1995). The BIC is given by

$$\text{BIC} = -2\log_e \mathcal{L}(\hat{\mathbf{m}}) + M\log_e N\,, \tag{1}$$

where M is the number of model parameters and N is the number of data. Since the BIC is based on the negative log likelihood, the model with the smallest BIC is selected as the preferred model. In comparison to the also commonly-used Akaike information criterion (AIC, Akaike (1973))

$$\text{AIC} = -2\log_e \mathcal{L}(\hat{\mathbf{m}}) + M\,, \tag{2}$$

the BIC corrects for the number of data and favors simpler models than the AIC for $N >$8. Since this study addresses large data sets, the BIC is used here for model selection.

WORK COMPLETED

In the second year of this project, the work was focused on extending the available inversion schemes to allow for quantitative model selection. This is considered to be the last step towards analyzing several

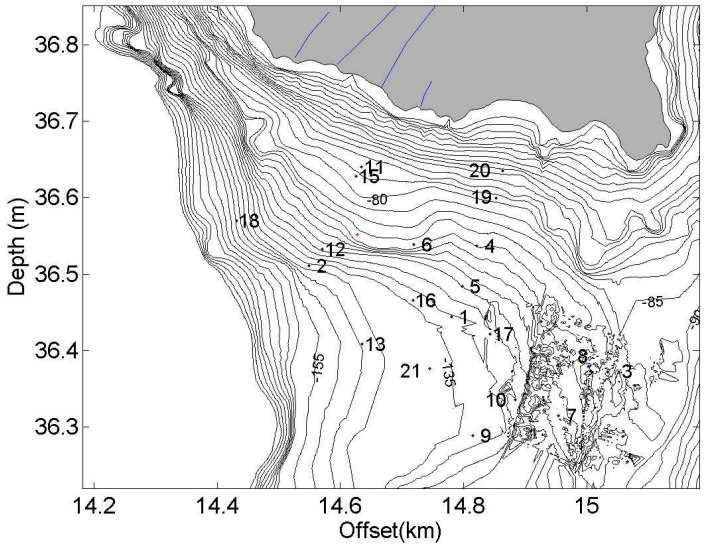

Figure** 1: **Experiment sites on the Malta Plateau.

sites for spacial sediment variability. The geoacoustic model selection inversion was applied to several data sets collected on the Malta Plateau (see Fig. 1). The procedure for data collected at Site 13 on the Malta Plateau (Fig. 1) is discussed in detail in Dettmer et al. (2008b) and results are shown below.

RESULTS

This section applies the inversion and model selection to seismo-acoustic reflection data collected April 6, 2002, during the Boundary02 experiment at 36° 24.515' N, 14° 38.142' E on the Malta Plateau, Mediterranean Sea. The acoustic data were generated with an electro-mechanical impulsive source (GeoAcoustics 5813B Geopulse boomer) with a short pulse length (<1 ms) and a broad bandwidth (0.5–10 kHz). Data were recorded at a single receiver that was part of a vertical line array of 4 hydrophones, and sampled at 48 kHz. The hydrophone used in this data set was at 124-m depth and the water depth was 144 m. The sound-velocity profile was fairly constant with the sound velocity varying less than 5 m/s over the water column. The source was towed at 0.3-m depth.

Figure 2 shows part of the seismo-acoustic traces (in reduced time) with the solid lines across traces indicating the part of the bottom response used to compute reflection coefficients. Figure 2 also illustrates the need for an objective model selection criterion. Reflected energy is most concentrated around the water-sediment interface reflection at 0.109 s and at a later event at about 0.114 s (times are given here for the shortest-range trace). In both instances, the events are spread out in time and the model parameterization is not obvious: both zones could contain reflections from one or more layers.

Reflection-coefficient data as a function of grazing angle and frequency were computed from time-windowed direct and bottom-reflected arrivals using the method of Holland (2003) and are shown in Fig. 3. In this case, the bottom response is time windowed to approximately 6-m depth below the seafloor, as indicated in Fig. 2. The data are averaged into 14 frequency bands from 1000–5300 Hz using a Gaussian frequency average (Dettmer et al. 2007) with a fractional bandwidth of 1/20 (this

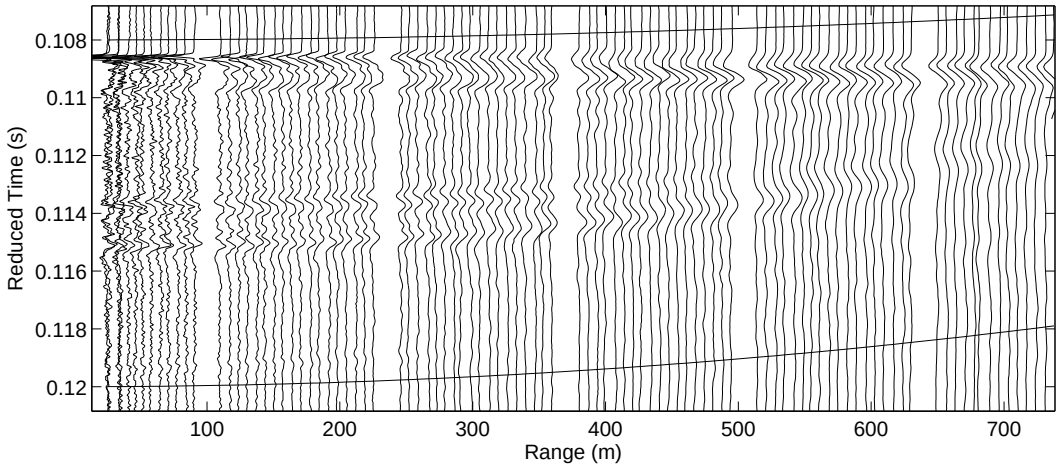

***Figure** 2: Seismo-acoustic traces (in reduced time, with 1512-m/s reducing velocity) collected at site 13 on the Malta Plateau.*

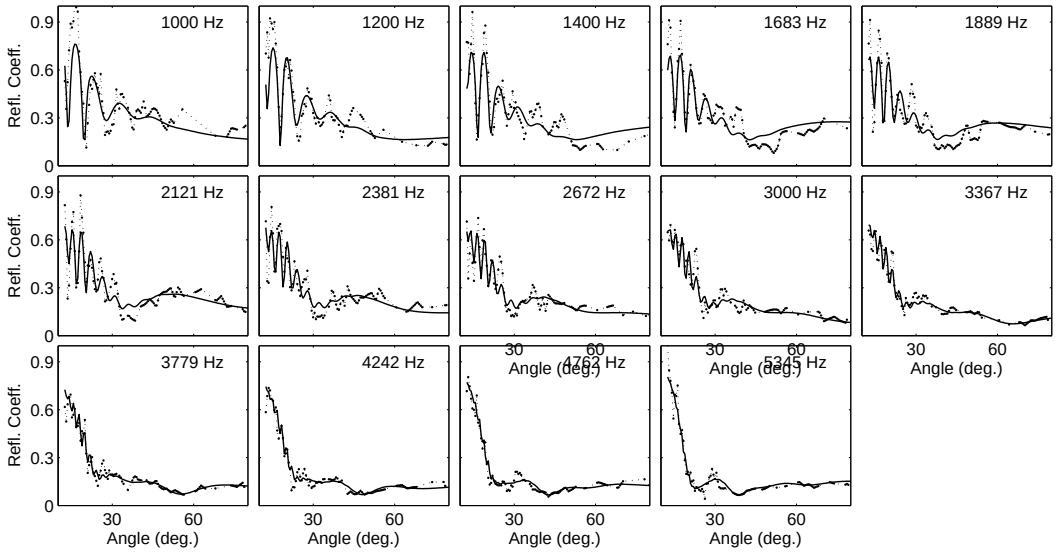

***Figure** 3: Reflection-coefficient data as a function of grazing angle and frequency (band centers indicated). The solid line indicates the best fit obtained from the MAP parameters of the model selected by the BIC (5 layers).*

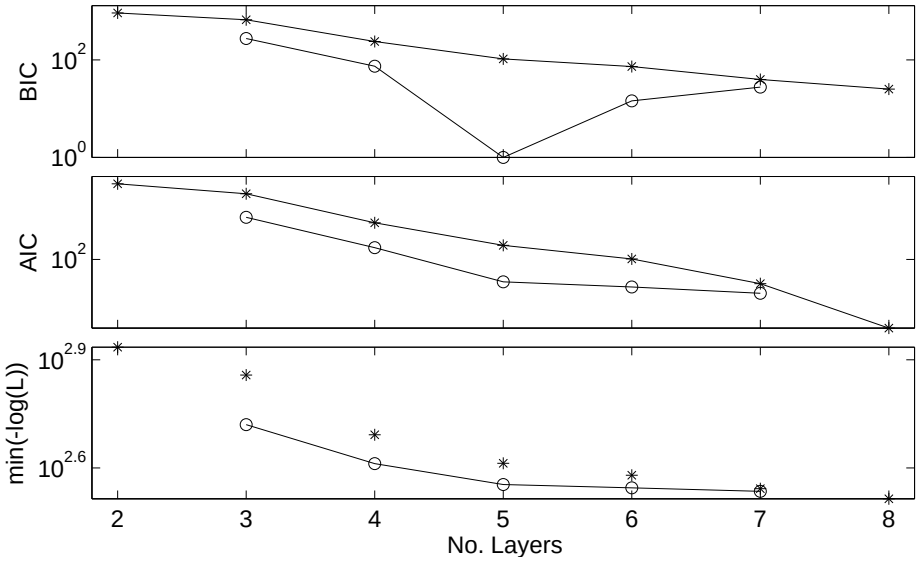

***Figure** 4: BIC, AIC and negative log likelihood for group A (open circles) and group B (asterisks) on a $\log_e$ scale. Note that for presentation purposes BIC and AIC values have been shifted so that the minimum value of each is unity.*

bandwidth was found to retain structure in the reflection coefficient data while reducing noise). The data are interpolated onto a uniform spacing in angle; points with a signal to noise ratio of less than 6 dB were excluded. Further, interpolated data that fall into recording gaps (due to experiment design) are excluded from the inversion. This results in approximately 90 data at each frequency.

The model selection study was carried out using two groups of models. Group A assumes an increasing number of layers for the uppermost part of the sediment (corresponding to the reflected arrivals beginning at 0.109 s in Fig. 2) and a single reflector at depth (corresponding to the deeper reflected arrivals at 0.114 s). Group B considers the same numbers of layers for the uppermost part of the sediment but contains two reflectors (i.e., a layer) at depth. This way both zones that show reflections in the time series are addressed by the model selection. Group A includes 5 models with 3 to 7 layers and group B includes 7 models with 2 to 8 layers.

To obtain inversion and model selection results, data-error statistics were estimated from the reflection-coefficient processing (Holland 2003) and optimization was carried out using plane-wave reflection-coefficient inversion (Holland et al. 2005). Prior bounds were chosen to be uniform over wide intervals. However, the thin topmost layers were differentiated by choosing prior bounds for layer thickness from 5–50 cm. This seems justified since the uppermost events in the time series (Fig. 2) extend over about 1.5 m (at 1500 m/s). These priors also allow the more complex models to exactly represent the simpler models. The resulting likelihood values were used to calculate the BIC for all models and results are shown in Fig. 4. All values are plotted on a $\log_e$ scale since the range of values is large (due to the fact that the 2- and 3-layer models are very unlikely with high BIC and AIC values). This figure shows that, based on the minimum value of the BIC, the 5-layer model from group A is selected. The BIC values for the models in group B do not reach values as low as those for group A, but consistently decrease with more complex parameterizations. This result indicates that the data show high sensitivity to the presence of several layers close to the sediment water interface. In addition, the simpler models of group A are preferred over the models of group B which contain more structure at

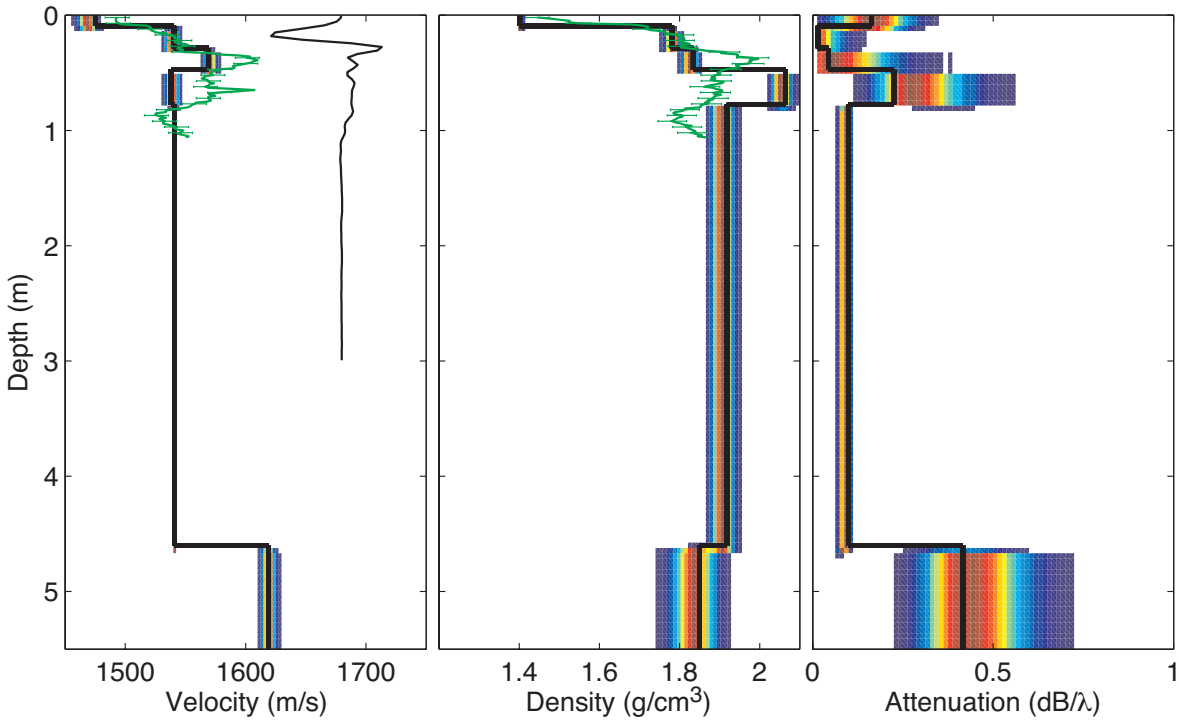

***Figure** 5: Marginal-probability depth distributions and MAP sediment profiles (solid black line) for the 5-layer model of group A (selected by BIC). A core (green) taken on site is shown for comparison, with error bars shown for every fifth datum. The left panel also shows the acoustic-source pulse (at 1500 m/s).*

depth. Figure 4 also shows the negative log likelihood (i.e., the data misfit) for all parameterizations. It is important to note that these values consistently decrease with an increasing number of layers. The BIC depends strongly on the likelihood values, and a better fit to the data for more complex models is important for the BIC to yield meaningful results. For comparison, the figure also shows the AIC values, with the minimum for both groups occurring at the models with most layers.

Once the preferred model parameterization was identified using the BIC, the data residuals for this model were used to compute a non-parametric estimate of the data error covariance matrix at each frequency (Dettmer et al. 2008a). Posterior statistical test were carried out for raw residuals $\mathbf{d} - \mathbf{d}(\hat{\mathbf{m}})$ and for standardized residuals $\mathbf{C}_d^{-1/2}[\mathbf{d} - \mathbf{d}(\hat{\mathbf{m}})]$ to ensure these estimate and the assumptions of random Gaussian errors were reasonable. The runs test for randomness failed at all 14 frequencies (at the 0.05 level) for raw residuals; however, standardized residuals passed the test at 13 out of 14 frequencies. The Kolmogorov-Smirnov test for Gaussianity was passed (at the 0.05 level) 4 times for the raw residuals and 9 times for the standardized residuals. Overall, the results of the statistical tests suggest the data error statistics are reasonably well quantified, providing confidence in the inversion results. The estimated error covariance matrices were then used in the integration of the PPD by MH sampling.

The integration was carried out for two models from group A, the 5-layer model that was picked due to the BIC and the 3-layer model to observe the variability in results with model selection. The fit to the measured data for the 5-layer model is shown in Fig. 3.

Figure 5 shows the MAP sediment profiles and associated uncertainties in terms of marginal-probability

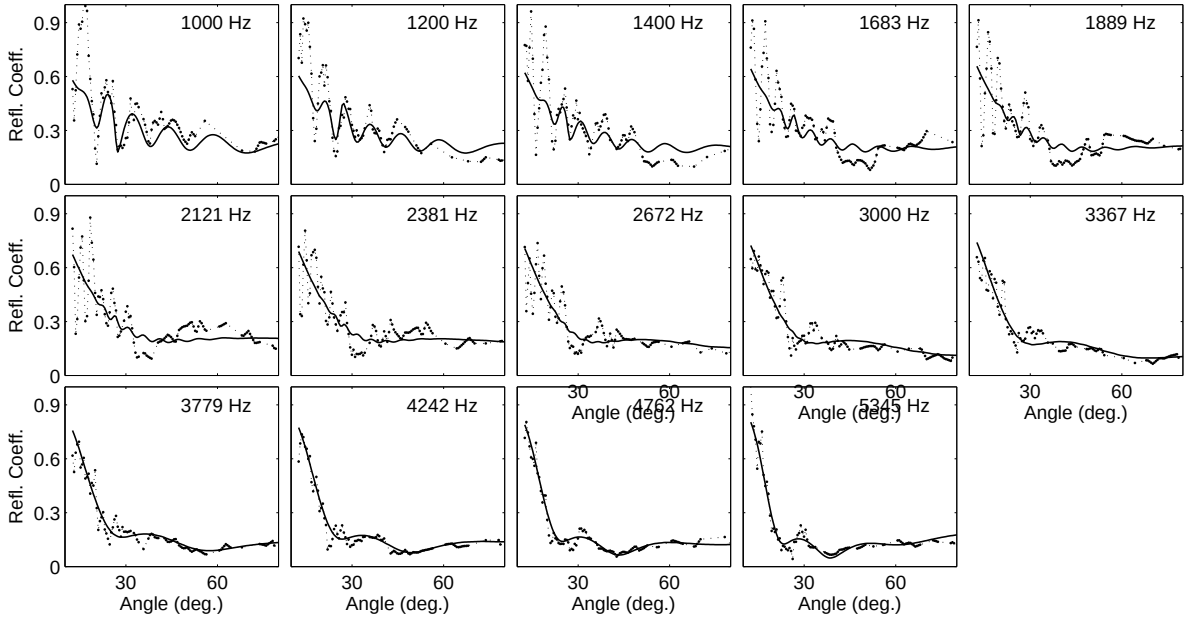

***Figure** 6: The best fit obtained from the MAP parameters of the 3-layer model of group A.*

depth distributions for the 5-layer model. The uncertainties are obtained from a large random subset of the PPD (4×10^5 models). The inversion results indicate 4 layers within the upper meter of the sediments. Figure 5 also shows the acoustic-source pulse (at 1500 m/s) for comparison with the inversion results: Note that the pulse length is large compared to the layered structure resolved in the upper sediments. The appearance and location of the thick layer is consistent with what would be expected from the time-domain data (Fig. 2) where no significant reflections occur between 0.11 s and 0.113 s. The inversion result also shows an interface at about 4.5-m depth. Confidence for the half-space parameter values, particularly density, is low, likely because the half-space lacks a lower reflector and hence data information is available only from the reflection off the upper interface.

Figure 5 also shows the sound velocity and density estimates from a shallow gravity core taken at the site. The core error bars, shown for every fifth datum, represent measurement errors associated with the time-of-flight and gamma-ray attenuation density estimates for a perfectly calibrated system, but do not include errors due to sediment disturbance from sampling, retrieval, calibration, and storage. Note that the core represents a highly-localized sample (10-cm diameter compared to the $\sim$100 m experiment footprint) and that spikes in the core values could be due to anomalies such as sea shells. The core indicates complicated structure in the upper part of the sediment. The inversion results show similar fine structure to the core. In particular, the density profile estimate from the inversion matches the core profile estimate well: the location of interfaces in the inversion result coincides with interfaces in the core. Below about 0.5-m depth, the inversion results appear to represent an average value in the sound velocity estimated by the core.

Figures 6 and 7 show inversion results for the 3-layer model from group A. The fit to the data (Fig. 6) is considerably worse than that for the 5-layer model (Fig. 3), and the resulting low log likelihood value dominates the BIC, resulting in the rejection of the 3-layer model. However, high frequencies show a much better fit than low frequencies, suggesting that the model is appropriate for the shallowest structure.

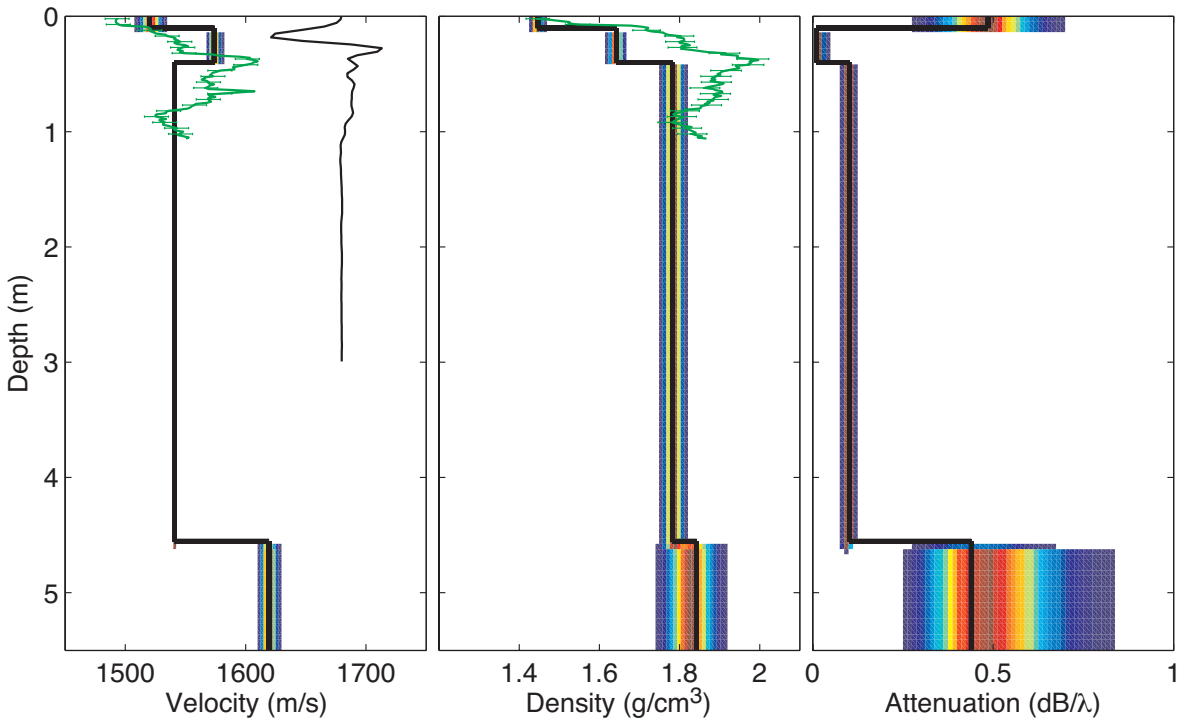

***Figure** 7: Marginal-probability depth distributions and MAP sediment profiles (solid black line) for the 3-layer model of group A. A core (green) taken on site is shown for comparison. Core error bars are shown for every fifth datum on the core.*

IMPACT/APPLICATIONS

The ability to remotely obtain (i.e., without direct sampling) seabed parameters has important implications for science (e.g., providing data for understanding sediment processes), the Navy (e.g., improving databases for ASW and MCM), as well as many commercial applications (e.,g., pipeline or cable laying). A particular strength of the present work is quantifying the uncertainties of the seabed parameters.

RELATED PROJECTS

Broadband Clutter JRP project (NURC, ARL-PSU, DRDC-A, NRL)
ONR QPE Uncertainty Program

REFERENCES

S. E. Dosso. Quantifying uncertainty in geoacoustic inversion. I. A fast Gibbs sampler approach. *J. Acoust. Soc. Am.*, 111:129–142, 2002.

J. Dettmer, S. E. Dosso, and C. W. Holland. Joint time/frequency-domain inversion of reflection data for seabed geoacoustic profiles. *J. Acoust. Soc. Am.*, 123:1306–1317, 2008a.

S. Chib. Marginal likelihood from the Gibbs output. *J. Am. Stat. Assoc.*, 90:1313–1321, 1995.

G. Schwartz. Estimating the dimension of a model. *Ann. Statist.*, 6:461–464, 1978.

R. E. Kass and A. E. Raftery. A reference Bayesian tets for nested hypotheses and its relationship to the schwarz criterion. *J. Am. Stat. Assoc.*, 90:928–934, 1995.

H. Akaike. *Proc. 2nd Int. Symp. Information Therory*, chapter Information theory and the extension of the maximum likelihood principle, pages 267–281. Akademiai Kiado, Budapest, 1973.

J. Dettmer, S. E. Dosso, and C. W. Holland. Model selection and Bayesian inference for high resolution seabed reflection inversion. *J. Acoust. Soc. Am.*, 2008b. submitted.

C. W. Holland. Seabed reflection measurement uncertainty. *J. Acoust. Soc. Am.*, 114:1861–1873, 2003.

J. Dettmer, S. E. Dosso, and C. W. Holland. Full wave-field reflection coefficient inversion. *J. Acoust. Soc. Am.*, 122:3327–3337, 2007.

C. W. Holland, J. Dettmer, and S. E. Dosso. Remote sensing of sediment density and velocity gradients in the transition layer. *J. Acoust. Soc. Am.*, 118:163–177, 2005.

PUBLICATIONS

J. Dettmer, S. E. Dosso, and C. W. Holland. Model selection and Bayesian inference for high resolution seabed reflection inversion. *J. Acoust. Soc. Am.*, 2008a. [submitted, refereed].

J. Dettmer, S. E. Dosso, and C. W. Holland. Joint time/frequency-domain inversion of reflection data for seabed geoacoustic profiles. *J. Acoust. Soc. Am.*, 123:1306–1317, 2008b [published, refereed].

J. Dettmer, S. E. Dosso, and C. W. Holland. Full wave-field reflection coefficient inversion. *J. Acoust. Soc. Am.*, 122:3327–3337, 2007b [published, refereed].

J. Dettmer, S. E. Dosso, and C. W. Holland. Uncertainty estimation in seismo-acoustic reflection travel-time inversion. *J. Acoust. Soc. Am.*, 122:161–176, 2007a. [published, refereed]